DOCUMENT RESUME

ED 449 397 CG 029 845

AUTHOR	Harwood, Henrick; Bazron, Barbara; Fountain, Douglas
TITLE	Performance Measurement for Substance Abuse Treatment Services. Integrated Evaluation Methods. Revised.
INSTITUTION	Caliber Associates, Fairfax, VA.; National Evaluation Data and Technical Assistance Center, Rockville, MD.; Lewin Group, Inc., Washington, DC.
SPONS AGENCY	Substance Abuse and Mental Health Services Administration (DHHS/PHS), Rockville, MD. Center for Substance Abuse Treatment.
PUB DATE	1999-07-00
NOTE	51p.; This document is being made available through Caliber/NEDS Contract No. 270-97-7016.
CONTRACT	270-94-0001
PUB TYPE	Reports - Descriptive (141)
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	*Activities; *Cost Effectiveness; *Evaluation Methods; *Outcomes of Treatment; *Substance Abuse

ABSTRACT

This paper presents state-of-the-art models addressing issues related to coordination of treatment and evaluation activities, and integration of clinical, performance, and evaluation information. Specifically, this concept paper contains a discussion of the need for and types of cost analyses for CSAT treatment evaluation and knowledge-generating activities. The theme of this paper is that there is a foundation for designing and undertaking comparisons of client outcomes across providers that make adjustments for the real differences in clients served by different providers. This foundation lies in undertaking case-mix adjustments when making such comparisons, rather than directly comparing the unadjusted results obtained for a given provider. Such information is useful to providers, as well as managed behavioral health organizations, policymakers, and clients and their friends and families. Appendix A is "Integrated Evaluation Methods Package: A Guide for Substance Abuse Treatment Knowledge-Generating Activities--Executive Summary" and Appendix B is "Editor's Note." (Contains 1 figure, 3 tables, and 19 references.) (MKA)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 449 397

INTEGRATED EVALUATION METHODS

PERFORMANCE MEASUREMENT FOR SUBSTANCE ABUSE TREATMENT SERVICES

CG029845

Revised July 1999

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☒ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

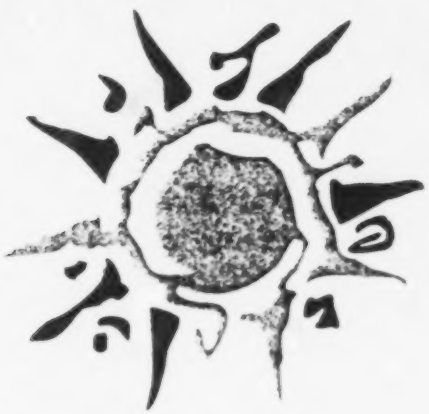
NEDTAC

CSAT
Center for Substance Abuse Treatment
SAMHSA

The Lewin Group

BEST COPY AVAILABLE

INTEGRATED EVALUATION METHODS



PERFORMANCE MEASUREMENT FOR SUBSTANCE ABUSE TREATMENT SERVICES

Prepared by:

Henrick Harwood
Barbara Bazron
Douglas Fountain

The Lewin Group
3130 Fairview Park Drive
Suite 800
Falls Church, VA 22042

Revised July 1999

This document was produced by the Center for Substance Abuse Treatment, Department of Health and Human Services, Caliber/NEDTAC Contract No. 270-94-0001 and is being made available through Caliber/NEDS Contract No. 270-97-7016.

CSAT
Center for Substance Abuse Treatment
SAMHSA

BEST COPY AVAILABLE

TABLE OF CONTENTS

Page

FOREWORD i

ACKNOWLEDGMENTS ii

I. INTRODUCTION 1

1. CONTEXT FOR THE PERFORMANCE MEASUREMENT PAPER 2

2 IMPORTANCE OF PERFORMANCE MEASUREMENT IN EVALUATIONS 3

II. WHY ASSESS PROVIDER PERFORMANCE? 8

III. APPROACH TO PROVIDER PERFORMANCE ANALYSIS 12

1. SELECT PERFORMANCE MEASURES: CLIENT OUTCOMES 15

2. SELECT MEASURES OF CLIENT SEVERITY PREDICTIVE OF FUTURE OUTCOMES 15

3. SELECT MEASURES OF TREATMENT SERVICE STRUCTURE AND DESIGN 16

4. STATISTICALLY ANALYZE RELATIONSHIP OF OUTCOMES TO SEVERITY MEASURES 17

5. GENERATE PREDICTED OUTCOMES FOR CLIENTS AND FOR PROVIDERS 17

6. COMPARE ACTUAL OUTCOMES TO PREDICTED (CASE-MIX-ADJUSTED) OUTCOMES 18

7. TEST STATISTICAL SIGNIFICANCE OF PERFORMANCE MEASURES . 18

IV. HISTORY OF PERFORMANCE MEASUREMENT IN THE SUBSTANCE ABUSE FIELD 20

1. OREGON OFFICE OF ALCOHOL AND DRUG ABUSE PROBLEMS 22

1.1 Data Collection and Performance Indicators 22
1.2 Analysis 24
1.3 Impact 25

TABLE OF CONTENTS (CONTINUED)

Page

2. METHADONE TREATMENT QUALITY ASSURANCE FEASIBILITY STUDY 26

V. SUMMARY **29**

VI. CONCLUSION **32**

REFERENCES **33**

APPENDIX A: INTEGRATED EVALUATION METHODS PACKAGE: A GUIDE FOR SUBSTANCE ABUSE TREATMENT KNOWLEDGE-GENERATING ACTIVITIES—EXECUTIVE SUMMARY **A-1**

APPENDIX B: EDITOR'S NOTE **B-1**

FOREWORD

Over the last 10 years the Center for Substance Abuse Treatment (CSAT) has accumulated a great deal of experience in substance abuse treatment evaluation implemented through coordinating centers, cross-site efforts, and national studies. The importance and value of integrating ongoing evaluation activity into a system for treating substance abuse problems is widely recognized. Also widely recognized, however, is that current evaluation-generated knowledge and practice are often under-utilized, due in part to the lack of an integrated approach to capturing information with which to measure and improve treatment effectiveness, efficiency, and performance. CSAT recognizes that such an integrated evaluation approach will more effectively support current and future knowledge generating activities.

Based on a decade of evaluation experience, CSAT has developed the Integrated Evaluation Methods (IEM) Package, a series of conceptual and methodological applications, including concept papers, technical assistance materials, and analytic tools, to enhance CSAT-funded evaluation activities. Products in the IEM Package are organized within an evaluation framework constructed on the basis of accumulated experiences among internationally known treatment service evaluation professionals. Thus, the framework is based upon evaluation strategies, structures and approaches appropriate for substance abuse treatment evaluators and providers. The framework follows a standard set of evaluation activities: planning, selecting a design, developing data requirements and collection instruments, collecting and analyzing the data, and reporting the evaluation findings. (A summary description of the IEM Package is contained in Appendix A to this document.)

This concept paper and its companion documents, *Integrated Evaluation Methods: A Guide for Substance Abuse Treatment Knowledge-Generating Activities; Self-Adjusting Treatment Evaluation Model; Adding "Value" to CSAT Demonstrations: The What, How and Why of Cost Analysis; Building Team Capability to Fully Implement and Utilize the Self-Adjusting Treatment Evaluation Model,* and *Client Levels of Functioning as a Component of Substance Abuse Treatment Services Evaluation* present state-of-the-art conceptual models addressing issues related to coordination of treatment and evaluation activities, and integration of clinical, performance and evaluation information. Specifically, this concept paper contains a discussion of the need for and types of cost analyses for CSAT treatment evaluation and knowledge-generating activities.

Sharon Bishop
Project Director
NEDTAC

ACKNOWLEDGMENTS

This paper, together with the companion documents listed in Appendix A, was developed for CSAT by the National Evaluation Data and Technical Assistance Center (NEDTAC) under the guidance and direction of Ron Smith, Ph.D., Program Evaluation Branch, Office of Evaluation, Scientific Analysis, and Synthesis (OESAS). Charlene Lewis, Ph.D., former Deputy Director, OESAS, supported this and other associated efforts, with the result that state-of-the-art evaluation concepts were incorporated into many of CSAT's and SAMHSA's evaluation initiatives. Jerry Jaffe, M.D., former Director, OESAS, also contributed his breadth of experience in the substance abuse treatment and evaluation fields and his dedication to high quality treatment services evaluation and provided the national level leadership necessary for CSAT to promulgate these activities.

Caliber Associates was the prime contractor for NEDTAC in partnership with Computech, Inc.; the Lewin Group; Capital Consulting Corporation; the Center for Substance Abuse Research (CESAR), University of Maryland; the Alcohol Research Group (ARG), Public Health Institute; the Drug Abuse Research Center (DARC), University of California, Los Angeles; and the Urban Institute. Many people within the NEDTAC team contributed to this effort. Henrick Harwood, Barbara J. Bazron, and Douglas Fountain of the Lewin Group were responsible for development and writing of the initial document. Patricia Devine, Jacquelyn Lowery, Melody Moore, and Judith Walton, Caliber Associates, were responsible for review, design, and editing of the final document. Thanks are also due to Robin Walthour, Erica Gordon Sorohan, and Donna Caudill for their many contributions to the quality of this document.

I. INTRODUCTION

The Center for Substance Abuse Treatment (CSAT) supports the integration of ongoing evaluation within substance abuse treatment activities so as to demonstrate treatment service effectiveness and to improve treatment services and their outcomes. To this end, CSAT recommends the use of state-of-the-art evaluation methods and tools in planning, designing, and implementing treatment services evaluations. This document provides a discussion of the increasing importance of provider performance measurement and analyses and an explanation of the case-mix adjustment methodology.

Performance measurement is a method for comparing performance with defined targets/benchmarks, the performance of one provider with other providers, or the performance of a provider with its own prior performance. Performance measurement is undertaken for the purpose of identifying who is meeting or exceeding expectations and who is not; identifying management, service, clinician, and client variables which contribute to outcome; and making adjustments in treatment service delivery in order to achieve desired outcomes.

Performance measurement methods can address the following types of questions:

- Which treatment providers are more effective? Which are less effective?
- Which treatment approaches (i.e., modalities, services, bundles of services, clinical approaches) are most effective? For which clients?
- Why are particular treatment approaches more or less effective?
- Why are particular providers more or less effective?

While the substance abuse treatment field has historically shied away from addressing these difficult questions, similar analyses are being vigorously conducted in the general health field. It has been proven feasible to formulate and execute a system of performance measurement that offers valid and reliable information regarding treatment outcomes. Further, questions related to performance measurement are of importance to service providers and the organizations responsible for funding or managing networks of providers. Policy makers are constantly asking for evidence about whether and how well substance abuse treatment works for what type of client. Each stakeholder group has an interest in assuring that the necessary services are provided to achieve the desired outcome.

1. CONTEXT FOR THE PERFORMANCE MEASUREMENT PAPER

CSAT's major evaluation goals are to: (1) increase knowledge about substance abuse treatment services; (2) improve treatment services by applying knowledge gained through knowledge development and application (KD&A) activities; (3) develop analytic methods and approaches for use in knowledge-generating activities; and (4) develop substance abuse treatment analysis databases. To meet these goals, CSAT has been sponsoring KD&A initiatives including activities that focus on homelessness, marijuana use and treatment, managed care, women and violence, and opioid treatment, as well as the replicability of exemplary treatment approaches (e.g., methamphetamine treatment) and the evaluation of best practices for targeted populations (e.g., exemplary adolescent treatment).

CSAT's evaluation experiences have reinforced the fact that substance abuse treatment evaluation involves a standard set of tasks that generally occur in the following order:

- **Planning the evaluation**, which includes setting the evaluation goals and objectives that determine the overall parameters of the evaluation
- **Selecting the evaluation design**, which sets forth the overall strategy for establishing the evaluation questions, measurement approach, and generalizability of findings
- **Developing the data requirements**, which flow from the evaluation questions and measures and include SDU, clinician, cost, and client data
- **Developing data collection instruments**, which are based on the data requirements and are developed or selected from a standard inventory of instrumentation
- **Collecting the data**, which includes the development of data management processes and tools including quality control procedures, and collecting the data
- **Analyzing the data**, which involves developing an analysis plan and conducting multiple levels of comparison; the analysis process is governed by the analysis plan and intended products and target audience(s)
- **Reporting the evaluation findings**, which includes evaluation knowledge dissemination and application within field.

CSAT has directed the development of evaluation concepts, methods, and tools to support these evaluation tasks. The evaluation tasks and corresponding evaluation methods are summarized in Exhibit I, Appendix A. As shown, performance measurement considerations are most

appropriately addressed in stage one in the evaluation process: planning the evaluation. A full discussion of the CSAT evaluation analytic framework and the other evaluation concepts and tools, is presented in the concept paper: *Integrated Evaluation Methods: A Guide for Substance Abuse Treatment Knowledge Generating Activities*. This document is fully referenced in Appendix A.

2. IMPORTANCE OF PERFORMANCE MEASUREMENT IN EVALUATIONS

Individual service providers need appropriate performance measurement tools to assess and monitor performance. Those responsible for funding or managing a system or network of service providers need performance measurement data to assess whether providers are delivering services that maximize client outcomes and value. Several state substance abuse agencies are currently undertaking the challenging task of measuring the performance of service providers.

Potential approaches to performance measurement are outlined in this paper. In addition, potential variables for assessing provider performance, the rationale for and type of approach that could be taken to case-mix adjustment in performing comparisons, and current analytic techniques being used in the technical literature for performance measurement are presented.

This paper is one of a series of evaluation products developed by CSAT. The goal of all of these products is to provide CSAT-funded treatment service providers, staff, and other interested parties with up-to-date and accessible information on evaluation that is particularly relevant to the evaluations of demonstrations being supported by CSAT, including the Knowledge Development and Application (KD&A) activities. A key emphasis of the approach to evaluation taken by CSAT is that evaluation ***is*** and ***should be*** an integral part of treatment services management and part of an ongoing process of assessment and systems/services improvement.

The paper begins by examining why it is important to assess provider performance. The ultimate goal of any performance measurement strategy is to:

- Generate data that can be used for results-oriented management
- Provide external accountability and, ultimately
- Improve the quality of the services offered.

The output from provider performance measurement efforts should equip individual providers with insights that allow them to identify areas of potential strength and weakness in order to develop strategies to correct and improve performance. Networks can use these data to assess the performance of the various providers, both for quality purposes and to decide which providers to maintain.

The next section discusses potential approaches to and issues in assessing provider performance measurement. A general methodological approach is presented and important issues in developing a provider ranking are identified and discussed. In addition, some of the most salient issues in interpreting and utilizing provider performance rankings are identified and reviewed.

The paper continues with a description of a potential case-mix methodology that could be applied to analysis of the CSAT initiatives. The CSAT initiatives—both the KD&A activities and the Cross-Site Demonstrations—present an opportunity to apply this methodology. Indeed, performance comparisons across providers may be among the most innovative analyses that can be undertaken for the CSAT initiatives. The evaluation design in the current initiatives explicitly includes the capability to make comparisons across the various types of providers. An essentially identical core of data is being collected for representative groups of clients that are being served by the respective provider sites. This will make it possible to:

- Examine and compare the unadjusted performance of the respective providers
- Compare performance after making a case-mix adjustment for the addiction severity of clients in each provider
- Assess the statistical significance of observed deviations
- Analyze the cost effectiveness and cost benefits of service delivery units.

In order to provide some context and history to the approach to modeling performance, this paper will review several provider performance measurement efforts in substance abuse treatment. It should be noted that within the substance abuse treatment field there have only been a few very important efforts at assessing/comparing provider performance. Some of these have resulted in publications and conference presentations (e.g., McLellan et al., 1994, Harwood et al., 1996, and Phibbs et al., 1997).

There are several initiatives currently under way that are attempting to assess provider performance at the state level. Oregon and Maine have both operated a system for several years. More recently, seven states have simultaneously and uniformly joined in the NIDA-sponsored pilot test of the Methadone Treatment Quality Assurance System. (A description of these efforts is also contained within this document.)

CSAT has an opportunity to make advancements in the field of performance measurement and analysis by applying these strategies to the treatment field. As the Federal government provides the states with major support through the Block Grant funding mechanism, it is logical for policy makers to ask about how well those funds are being spent. The CSAT initiatives represent a major step toward answering those questions.

One of the core objectives—and design features—of these initiatives is to acquire consistent data across the participating treatment service delivery units (SDUs). This will ultimately allow CSAT to make direct comparisons across treatment categories in terms of average outcomes of clients served by different types of SDUs. However, if these comparisons are to be meaningful, they must go beyond a process that involves making simplistic, unadjusted comparisons of outcomes of clients served by different providers.

In summary, the theme of this paper is that there is a scientific foundation for designing and undertaking comparisons of client outcomes across providers that make adjustments for the real differences in clients served by different providers. This foundation lies in undertaking case-mix adjustments when making such comparisons, rather than directly comparing the unadjusted results obtained for a given provider.

Case-mix adjustment is a methodology which controls for systematic, measured differences in the nature and severity of clients served across providers. It involves relatively straightforward application of standard multivariate analytical techniques. This methodology provides a means of accounting for the wide range of different outcome objectives of providers.

The following examples illustrate the use of case-mix adjustment in a variety of situations:

- For a provider that delivers brief interventions to a specific group of clients (e.g., women with children) case-mix adjustments would be used to compare the different SDUs because clients may differ from SDU to SDU on demographic and socio-economic characteristics

- For a treatment services provider that includes a number of different outpatient SDUs in six different locations, case-mix adjustments would be used to address the difficulty in drawing conclusions about the relative success of providers in helping clients access such services based on differences between inner-city and rural providers

- For treatment services that involve 6- or 12-month treatments, case-mix adjustments would be used to better assess whether there are different outcomes for the different planned treatment durations.

The fundamental question is, "Which projects, treatment approaches, and/or models are more effective in improving client outcomes?" It is anticipated that strong evaluation results might yield information about which models of care should be singled out for further development and study, and ultimately for wider dissemination and adoption within the substance abuse treatment field. The methodology proposed in this paper can help ensure the validity of the evaluations conducted, lending strength and credibility to evaluation findings.

As stated, the focus of this paper is performance measurement. In order to demonstrate the value of their services, providers must have some type of performance measurement system or method for comparing performance of defined targets/benchmarks with its own prior performance. To ensure meaningful comparisons of performance, client outcomes data must be adjusted to reflect differences in clients served.

It is important to note the link between this paper and the document *Adding "Value" to CSAT Demonstrations: The What, How, and Why of Cost Analysis.* (See Exhibit I, Appendix A.) In the current value-driven environment, outcomes data alone are not sufficient performance measures. Value is assessed in terms of outcomes relative to costs. This paper discusses the need for provider performance assessment and a methodology for adjusting for client differences (i.e., case-mix adjustment). This paper focuses on a methodology for assessing the outcome component; the cost paper focuses on methodology for assessing value in terms of cost-effectiveness and cost-benefit. The performance measurement discussion is presented in six sections:

- Section I provides an introduction

- Section II explains the need for provider performance assessment—providers need such technology in order to understand how well they are serving their clients, to identify areas of strength as well as weakness, and to devise strategies to improve services

- Section III presents a six-step multi-provider approach to provider performance analysis
- Section IV details the history of performance measurement in the substance abuse field and discusses several recent performance measurement initiatives in the substance abuse field
- Section V presents a summary of the paper
- The conclusion is presented in Section VI.

Two appendices provide additional information about CSAT evaluation concepts and their interrelationships.

II. WHY ASSESS PROVIDER PERFORMANCE?

Performance measurement technology is critically needed to answer the following types of questions:

- How well do my treatment services work (a provider assessing its own treatment)?
- Which treatments and providers are better (a buyer of services trying to buy wisely)?
- Is the treatment system identifying the best and poorest providers (policy makers)?

These are the most fundamental questions posed regarding treatment in budget and policy discussions. Providers need performance measurement technology in order to understand how well they are serving their clients, to identify areas of strength as well as weakness, and to devise strategies to improve services. Entities that pay for services (state and local government agencies, private health plans, and managed care systems) all need information about how well particular providers are doing. Indeed, Congress and the Substance Abuse and Mental Health Services Administration (SAMHSA) have a continual interest in knowing what treatment approaches work best and how well treatment works for particular types of clients. Answers to these questions are also important for individuals and families searching for help with substance abuse problems, and for operators of treatment systems and health/insurance networks.

Performance assessments are extremely useful to providers who want to implement the Self-Adjusting Treatment Evaluation Model. According to this model, the data derived from program evaluation can and should be used to inform decisions about what, if any, corrections should be made to service delivery in order to improve performance. These data can also be used to provide feedback to staff regarding those aspects of service delivery that contribute to desired client outcomes.

Rigorous analyses of performance data based on meaningful comparisons can be a strong management tool for quality assessment and improvement. An operating provider performance monitoring system could be designed to assess client outcomes for particular providers across a number of dimensions, including:

- Client use of alcohol and/or drugs
- Re-arrest/re-incarceration experience
- Employment

- Physical health and utilization of health care services
- Mental health and utilization of treatment service.

The same dimensions of performance that an individual provider would want to assess and monitor would be examined by a buyer of services, such as a state or local government, an insurance plan, and/or a managed behavioral health organization (MBHO). The challenge of the latter will be to differentiate between the effectiveness of the providers and the impact of the managed care arrangements. However, it may be equally if not more important to make comparisons across managed care plans, to the extent that clients may be in different plans. CSAT recently has undertaken the challenge through its Managed Care KD&A . Although until recently providers have not been ranked according to their performance on specific indicators, the pressures for increased accountability of providers and for efficiency by private and public systems and networks have created powerful incentives that are moving the field forward.

Policy makers want to know how treatment can be improved. More fundamentally, they want to know that the field itself can identify providers that are exemplary in treating clients, and providers whose performance is not adequate. Policy makers want accountability for the several billions of Federal and state funds put into the national treatment system. Often they have particular concerns about how to improve treatment for specific client populations that present different severity problems, patterns of drug use, etc. Performance measurement offers a set of methods and tools that can be systematized and used to routinely monitor and address the concerns of this critical constituency group.

More specifically, performance measurement can be applied to evaluate providers that receive support from SAMHSA Block Grants. While KD&As and Cross-Sites apply to several dozen providers, there are seven to eight thousand providers that get Federal support through individual states. There is a critical need to assess performance at the level of community providers in order to identify exemplary providers (and what makes them exemplary) and providers that appear to do poorly. It is not possible to make such determinations rigorously without performance measurement and case-mix adjustment. Whether working with CSAT-funded treatment providers or the providers funded by states attempting to assess provider performance, there is an opportunity to devise a systematic approach to undertaking comparisons across providers. There is an enormous opportunity to learn from and improve the system of community-based providers with development and judicious application of a performance measurement system.

The traditional approach to assessment and assurance of provider quality has been all or nothing: providers were either licensed or accredited to provide care, or they were not. Licensure and accreditation of providers have been driven by a combination of standards and checks on credentialing of staff and practitioners, checks on procedures and record keeping about client assessment, treatment plans and progress, and checks on samples of client files and records to establish compliance with accepted standards of practice.

The relatively late arrival of provider performance measurement in the substance abuse treatment field as well as in the general health arena is due in part to provider concern that provider comparisons (particularly those not using case-mix adjustment) will single out those providers that tend to serve the most severely affected clients and populations, who will therefore exhibit the poorest average outcomes for their clients. The case-mix adjustment approach outlined in this document is intended to address exactly this problem and adjust effectiveness measures for client severity—level the playing field. To avoid masking important population and/or treatment service characteristics, however, one should always present both unadjusted and adjusted rates.

Another issue is the fact that the most meaningful outcome data is collected post-treatment, which means that the data can be very difficult and expensive to obtain, due to the challenge of performing follow-up studies with severe and/or chronic substance abusers. Major problems with non-response fundamentally weakens any findings and conclusions from such analyses. The manual *Staying in Touch: A Fieldwork Manual of Tracking Procedures* (see Appendix A), contains a discussion of methods for performing effective follow-up surveys of treatment clients.

Concerns about measuring provider performance are appropriate and well-founded. However, this does not obviate the importance and need for doing so. It does increase the importance of developing methodologies for data acquisition and analysis that can yield meaningful provider performance measures, as well as create an appropriate understanding of their limitations and implications.

The technology for performing provider performance assessment has developed substantially over the past 12 years. The advancement of electronic data systems may improve the feasibility and affordability of doing provider performance assessments. A great deal of data can potentially be accessed from administrative data files of treatment providers, payers for substance abuse treatment and other health care services (government agencies and/or health

plans and insurance companies), criminal justice authorities, social service, and social welfare agencies. The challenge will be to develop strategies to access such data and to develop approaches for using it in appropriate and meaningful ways.

III. Approach to Provider Performance Analysis

The use of performance measurement is rapidly increasing, driven by need and advancements in the technology. The implication of the prior sections is that provider performance analysis can be designed and implemented in such a manner as to adjust estimates for the severity and complexity of clients served by respective providers. Given that provider performance will be increasingly examined, and probably utilized by purchasers of substance abuse treatment, the challenge is to develop case-mix adjustment methodologies—particularly tools to assess the prognosis and needs of clients—that will "level the playing field" for providers that serve the most severely affected populations.

The general approach to quantitative assessment of provider performance is to compare the service provider's client outcomes with some standard. This involves a fairly straightforward process:

- Select outcome measure(s) for assessment
- Acquire valid and reliable data on clients of the provider and on comparison group(s)
- Calculate measure(s) for provider and comparison(s) standards
- Determine whether the provider has met standards, exceeded, or under-performed with respect to the standards.

This type of process is one which is generally performed for a system of providers. That is, an individual provider that wants to use provider performance tools to assess its own treatment services would generally be better off participating in a multi-provider assessment system rather than attempting to develop and operate their own assessment. There are several main reasons for this:

- Efficiencies in development and operation of the system in terms of selection/ development of client measures and instruments
- Efficiency in data collection/acquisition
- A stronger statistical basis for performing comparisons.

It is still possible for an individual provider to conduct its own performance assessment. The approach would involve the provider using historical data on client outcomes (cohorts treated in prior years) as the comparison for current clients. The obvious attraction of this strategy is that a

provider may believe that they serve a similar mixture of clients from one year to the next, justifying comparisons of outcomes of recent/current clients to former clients.

Individual providers that consider undertaking performance assessments should be cautious about using cohorts of former clients as the comparison population. There are both systemic and statistical reasons that former client cohorts might not be the best comparisons. Instead, providers should participate in multi-provider evaluations, where the comparison data base should have thousands of clients.

The systemic reason is that for a given treatment provider both client populations and treatment approaches change constantly and rapidly. Changes are being driven by new and evolving patterns of alcohol and drug use, changes in who is paying for treatment and how it is paid for, and for what services. As a result of new or changing payer priorities, a given provider may add or lose specialized services for women, adolescents, or particular ethnic groups that can change the mix of clients served. The funding agency may cut back on the duration or intensity of treatment they will pay for and therefore on the services the different client cohorts receive. These changes happen with increasing frequency. The impact of such changes reduces the utility and feasibility of using old client cohorts for comparison purposes. All of these are reasons to participate in a multi-provider evaluation with solid data about both clients and treatment services and costs of treatment.

Statistical reasons also mitigate against using former client cohorts. Former client cohorts may only offer hundreds of clients for comparison purposes, while it should be possible to acquire much larger numbers when pooling comparison clients across multiple providers. Although it will be possible to use the analytical statistical procedures outlined below to compare current clients to prior cohorts, a large number of clients in the comparison population is more desirable for valid statistical assessments of performance. A larger number of clients in the comparison population provides a higher level of ability to detect whether deviations between provider actual and expected outcomes are statistically significant, and are therefore "real" differences rather than inevitable random fluctuation. For example, the Criminal Justice Network Demonstration of CSAT may be primarily limited by the small number of treatment providers, and the fact that these programs serve two distinct populations: women and adolescents. While such small numbers of distinct providers do not prevent developing provider performance comparisons, they do dictate that there are few "degrees of freedom" to be associated with differences in expected and actual program performance. Hence, it will be difficult to tell if the observed differences are due to variations in local conditions, systemic

differences in treatment services delivered, or random chance. This is particularly important if the analysis involves an attempt to adjust for differences in client case-mix using statistical analyses.

This final point serves to emphasize the importance of strong design of an evaluation data base, with particular emphasis on the number of clients as well as the number of different providers that are represented in the evaluation data base. Appropriate planning and design will allow sample sizes to be selected based on statistical criteria. This will give the evaluator the best chance to derive meaningful and generalizable conclusions from their data collection and analysis activity given budget and time constraints.

The particular challenges and needs in developing case-mix adjusted provider performance efforts are best understood within the methodological framework which has evolved over the past 12 years. The general approach to analysis of provider performance for a given group of providers can be summarized in the following steps:

- Select client outcome measures believed to respond to quality of treatment
- Select measures of client severity that are predictive of responsiveness to treatment
- Select measures of treatment service structure and design, such as nature, amount, intensity, and cost of services actually delivered to clients
- Statistically analyze the relationship of outcomes to client severity measures and measures of treatment service structure and design
- Generate predicted outcomes for individual clients and sum for each provider
- Compare aggregated actual outcomes to predicted outcomes for clients of each provider
- Statistically test for significance of differences between expected and actual outcomes.

These are the steps used in studies which attempt to use case-mix adjustments to compare the performance of providers, whether they are substance abuse providers or hospitals. The following discussion attempts to further illuminate the process by identifying and discussing major issues involved in attempting to make a case-mix provider performance system operational for substance abuse treatment providers.

The primary message to keep in mind is that all of these steps need to be performed consistently for each provider—and for their sampled clients—being included in the performance measurement analysis. That is, a comprehensive data collection system must be designed with this application in mind, and implemented rigorously and consistently. This is absolutely necessary in order to allow rigorous comparisons to be made, based on appropriate case-mix adjustment methodology.

1. SELECT PERFORMANCE MEASURES: CLIENT OUTCOMES

The first requirement for developing a system to assess provider performance is to select specific aspects of client outcomes that are both meaningful in terms of understanding client progress, and can be measured with validity and reliability. It is likely that substance abuse treatment provider performance would be assessed based on a battery of measures of client outcomes, due to the complexity of substance abuse problems, and the multiple objectives providers must usually address. In fact, individual providers need to assess how well they function on multiple levels of performance. The two examples of provider performance systems discussed later in this document—the Oregon system and the Methadone Treatment Quality Assessment System—both examine multiple outcome dimensions (see discussion of these systems below).

The major design issues concern what, when, and how outcome data will be collected. Standardized instruments offer obvious advantages. There are many ways to measure outcomes—which are beyond the scope of the current report—but the most critical requirement is probably to select a measure or set of measures and to implement them rigorously and consistently.

2. SELECT MEASURES OF CLIENT SEVERITY PREDICTIVE OF FUTURE OUTCOMES

Often it is felt that assessing the range of problems typically presented by substance abuse clients will yield both the most accurate diagnostic picture, as well as the greatest predictive power for the same outcomes. This is the basic logic underlying research strategies that collect very similar behavioral and status data at intake to treatment and follow-up, such as use of the addiction severity data.

Some of the measures that might be reviewed as a part of a performance measurement system are usually substance use, employment status, health and mental health status, and involvement with the criminal justice system. Also, demographic (i.e., age, gender, ethnic identity) and socioeconomic characteristics (i.e., education, employment experience) are important predictors of treatment outcomes. The *Minimum Evaluation Data Set* document is a CSAT-developed tool for developing a uniform set of variables and response categories that may be helpful for this purpose. (See Appendix A.)

There is a great deal of effort to attempt to assess additional factors about clients that might differentially predict their outcome of treatment, or their likely response to different types of treatment. For example, the American Society of Addiction Medicine (ASAM) has developed "treatment placement" criteria designed to indicate appropriate levels of initial care (i.e., hospital inpatient, residential, day treatment, or standard outpatient care) based on symptoms presented by clients at intake. Managed care organizations also have their own placement criteria. There are also scales and systems attempting to assess client "readiness for treatment" such as those proposed by DeLeon, and Prochaska and DiClementi. These scales and systems, including the ASAM criteria, should be considered areas for additional development, since they have all been developed based on best practices or "expert judgment" and have only been the focus of research and development on a limited basis to date.

3. SELECT MEASURES OF TREATMENT SERVICE STRUCTURE AND DESIGN

The services available to and delivered to clients should be important predictors of client outcomes. In this stage of the effort it is important to define and accurately measure the services. These measures will be combined along with individual client characteristics to analyze client outcomes. It is important at this stage to differentiate between availability of types of services and whether or not clients utilized the services. The cost of treatment services should also be assessed at this stage. Cost can be incorporated into the modeling of client outcomes in the following stage of the analysis.

Both services delivered, and their costs are quite important to assess because it is critical to examine how the nature, duration, intensity, and expense of treatment services relate to client outcomes. In general, more intensive/expensive services should only be delivered when they contribute to improved client outcomes. It, however, raises a more complex question of, "How much more effective must a service be?" in order to justify its use if it is more expensive.

4. STATISTICALLY ANALYZE RELATIONSHIP OF OUTCOMES TO SEVERITY MEASURES

This step involves analyzing the ability to predict client outcomes (in Section I, above) using client severity measures (in Section II, above). This critical phase of the analysis identifies the types of clients that on average have better or worse outcomes. Subsequently, this will make it possible to identify whether the clients served by a provider in a particular period of time were average, more difficult, or easier to treat than clients of other providers or other study periods.

A similar analysis is performed for each of the outcomes assessed in the study or system. Thus, one analysis would look at the factors predicting client use of alcohol at follow-up, another would look at factors predicting client use of cocaine at follow-up, and a third might look at factors predicting client re-arrest at follow-up. This analysis typically uses multivariate techniques such as multiple regression and logistical regression, depending on the particular outcome measure. Outcome measures which are "yes" or "no," or categorical (i.e., relapse, return to treatment, re-arrest, employment) are modeled using logistical regression, while outcome measures that are continuous (i.e., earnings, health care expenditures) can be modeled with multiple regression.

The objective in this phase of the analysis is to identify severity/predictor measures that are strongly related to/predictive of each of the outcomes of concern in the performance assessment. In other words, it is at this stage that it is learned which factors appear to indicate which clients are more and less likely to have favorable outcomes, and to derive formulas that can be used to predict "expected" outcomes for individual clients (actually the rate expected for a group of clients with the given characteristics of the individual in question), and for providers (actually the rate expected for the population with the mix of characteristics served by the provider).

5. GENERATE PREDICTED OUTCOMES FOR CLIENTS AND FOR PROVIDERS

The analyses executed in the prior stage can now be used to develop predicted or expected outcomes for individual clients on each of the outcome dimensions that have been analyzed. This is done by taking the results of the respective prior logistical and/or multiple regressions, and inserting the predictive values for individual clients, which generates an average or expected outcome for a cohort of clients that all have that individual's given set of characteristics. Most commercially available statistical packages (e.g., Statistical Analysis

System, or SAS) have routines that automatically generate, and output or save such expected values for individual observations for use in further analyses. In addition, it is possible to calculate the expected outcome for the population served by a given provider by simply averaging the expected values calculated for each individual served by that provider.

If the outcome of concern were 6-month relapse to alcohol abuse, the prior stage logistical regression would predict the probability of alcohol relapse for each person in the sample based on their individual characteristics, and then by summing these probabilities for all of the clients of a provider one can calculate the expected relapse experience of an actual population. This is actually the case-mix adjusted expected outcome rate for the provider, since it yields an outcome rate that accounts for the composition of their clientele.

6. COMPARE ACTUAL OUTCOMES TO PREDICTED (CASE-MIX ADJUSTED) OUTCOMES

The test of a provider's relative performance is whether the actual outcome of his clients is better or worse than the case-mix adjusted, expected outcomes calculated in the prior stage. It is expected that providers with on average more severe clients would have higher rates of expected relapse than providers with less severe clients. Then, if actual 6-month relapse for a provider is 40 percent, and expected 6-month relapse is 45 percent, the provider is apparently performing somewhat better than would be expected given their mix of clients.

A similar comparison would be developed on every outcome measure for each provider being assessed in the performance system. If the various outcome measures have been selected carefully, looking at the entire set of comparisons for a given provider should give indications whether that provider is relatively more or less effective at addressing particular needs of clients, or has achieved a consistent level of accomplishment with multiple needs. It should identify whether the provider generally does better or worse than average—and by how much—on each outcome measure. Thus, a provider might learn that it does relatively well in reducing client alcohol and cocaine use at follow-up, but that their clients are not doing as well as expected on addressing educational deficits, or at getting jobs.

7. TEST STATISTICAL SIGNIFICANCE OF PERFORMANCE MEASURES

The science of assessing provider performance dictates that statistical criteria should be applied to determine whether a provider's apparent performance (being better or worse than

expected given client case-mix) could be due to chance. Simply due to random chance, the apparent performance of a provider can vary even if the quality of care and client case-mix has not changed. Thus, differences between actual and expected outcomes are usually tested to see if the deviation is more or less than two standard deviations. It is usually only providers diverging more than two standard deviations that are singled out and reported as statistically significant, and represented as posing cause for note (favorable or unfavorable). In practice, about 95 percent of providers will fall within two standard deviations of the mean (either above, or below). Thus very few—less than one in twenty—would be statistically deemed as either having exemplary (or poor) performance.

In the spirit of caution, most producers of provider performance studies advise that they be used very carefully. ***It is usually recommended that "outliers" with actual performance rates more than two standard deviations away from expected rates should be subjected to expert on-site assessments.*** Other providers with performance that is better or worse than expected (but the deviation is not statistically significant) should be advised of rankings, but not subjected to interventions. Another version of applying caution would be to assess providers over multiple reporting periods, again, to see if deviations are transitory (and therefore more likely due to chance) or lasting—and probably representing a more systematic set of patterns, be they related to provider practices or client characteristics not measured and incorporated in the case-mix adjustment methodology.

IV. History of Performance Measurement in the Substance Abuse Field

Several recent initiatives illuminate how the substance abuse treatment field is addressing issues of client outcome in assessing performance of providers and even of carve-out managed care plans. A number of states are incorporating the assessment of individual provider performance into the contracting and management of their systems. There has been some published research describing specific performance measurement systems and the current status of activity within states.

The nationwide Federal drug treatment system created and monitored in the early 1970's under the Special Action Office of Drug Abuse Prevention (SAODAP) in the White House required providers to collect and report data about clients' continued drug use while in treatment (assessed via random urine tests) and the average length of stay in treatment. At that time, greater length of stay was considered to be an indicator of better performance, since a number of studies had demonstrated significantly better outcomes for clients with greater lengths of stay. Providers were required to make regular data submissions about client admissions and discharges, as well as file periodic reports about the status of clients currently enrolled in treatment (particularly regarding recent drug test results).

This data system was known as the Client-Oriented Data Acquisition Process (CODAP), and was turned over to the respective states as they assumed responsibility for managing and funding the drug treatment system in the late 1970s and early 1980s. The legacy of CODAP operates in most states across the nation (data on client admissions and discharges), and a national system is currently being reconstructed (Treatment Episode Data Set or TEDS: Office of Applied Studies, 1995). The CODAP system was used under SAODAP to monitor provider performance, and continues to offer some capacity to do this in many states from data reported about client status at discharge.

There have been several publications and/or conference presentations on provider performance assessments. McLellan et al. (1994) attempted to answer the question "Are Some Substance Abuse Treatment Providers More Effective than Others?" using the Addiction Severity Index (ASI) on samples of clients treated at two different providers. While this study did not directly perform case-mix adjustment, the authors undertook to carefully examine the characteristics and status of clients in the two programs at both intake and 6-month follow-up. The conclusion was consistent with the thesis of this report: in order to compare provider

performance using client outcomes at follow-up it is necessary to examine and control for client characteristics at intake.

Phibbs et al. (1997) analyzed re-admission to substance abuse treatment in Veterans' Affairs medical centers across the nation. That analysis did directly perform case-mix analysis and adjustment, concluding that it is both feasible to perform and that client characteristics are major predictors of their re-admission rates. Failure to adjust for client characteristics would significantly bias conclusions regarding relative re-admission rates of respective medical centers. Harwood et al. (1996) have analyzed the Treatment Research Institute's National ASI database (compiled by the Treatment Research Institute) including data on clients from 59 providers for which intake and 6-month follow-up ASI's have been collected. This analysis found that ASI intake measures are strong predictors of client follow-up ASI scores, and demonstrated graphically how provider performance rankings using client outcomes are affected by adjustment for client characteristics.

At the present time there are several notable provider performance assessments under way. The National Institute on Drug Abuse (NIDA) is developing the Methadone Quality Assessment System, the most advanced effort to date. Also, various state substance abuse agencies are doing or developing provider performance assessments. The states of Oregon and Maine both use provider assessment in oversight and contracting with substance abuse treatment providers. The Oregon provider performance contracting system is built directly on the data system that has evolved from CODAP, as is the system operated by Maine. The Maine system is largely modeled after the Oregon approach, therefore only the Oregon system is described below.

Another notable recent development is the emphasis placed on provider performance assessment in the accreditation standards for managed behavioral health organizations (MBHO) proposed by the National Committee for Quality Assurance (NCQA). This accreditation represents an attempt to improve the accountability and credibility of MBHOs and of the substance abuse (and mental health) treatment which they provide to their beneficiaries. A major part of these standards is data-based reporting by MBHOs about service delivery, access to care, effectiveness of care, and client satisfaction. New provider performance requirements, however, raise the question of whether already limited funds will be diverted from actual client treatment services.

1. OREGON OFFICE OF ALCOHOL AND DRUG ABUSE PROBLEMS

Oregon sustains and regulates over 200 publicly funded substance abuse treatment providers, which serve over 100,000 clients per biennium. The Oregon Office of Alcohol and Drug Abuse Programs (OADAP) contracts with providers on a slot basis and pays a flat, statewide rate for each slot. For example, OADAP contracts with outpatient treatment providers at $2,058 per contracted slot.

Oregon has had a performance measurement and contracting system since 1994 (personal communication, Calvin Phillips, 1996). Oregon's system relies on standardized admission and discharge data collected on clients utilizing the following types of services: emergency non-hospital detoxification; two levels of residential treatment for adults; specialized residential treatment for women, pregnant women, and youth; and outpatient, including methadone maintenance (Oregon Office of Alcohol and Drug Abuse Programs, 1996). All publicly funded providers must supply data and comply with this system, including Oregon Health Plan (OHP, or Medicaid managed care) providers (Oregon Office of Medical Assistance Programs, 1996).

1.1 Data Collection and Performance Indicators

Admission and discharge data have been collected from drug and alcohol clients using the Client Process Monitoring System (CPMS) since 1982. All publicly-funded contractors, including OHP-funded providers, must submit data to the performance measurement system. Data are supplied in hard copy to OADAP, which then scans/enters the data.

Discharge data and comparisons between discharge and intake data are used to measure performance. Specific measures of performance have been established, and detailed definitions of each measure (including how it is calculated) are available. Specific measures used in the system are presented below. Performance indicators for most adult treatment services include:

- Change in employability
- Employment improvement
- Educational advancement/participated in self-help
- Not arrested during treatment
- Abstinent/drug free

- Completed treatment
- Referral to self-help
- Referral to alcohol and drug treatment.

Youth-specific performance indicators include:

- Educational advancement
- Participated in self-help
- Not arrested during treatment
- Abstinent at termination
- Completed treatment
- Academic school improvement
- Improved school attendance
- Behavior in school improved.

Women-specific performance indicators include:

- Employment maintained (full/part time)
- Employment status improved
- Progressed in school or training
- Participated in self-help
- Not arrested
- Abstinent
- Completed treatment
- Reduced use

- Complied with Children's Services Division agreement
- Mother abstinent 30 days before delivery.

Documentation does not indicate whether the data are compiled by staff through an interview process or whether CPMS relies on notes. Assessments of the validity and reliability of the data are not reported.

In addition to performance data, OADAP collects the following types of quantitative data and uses this data to establish contracts with providers. Providers are also compared using this data. The data collected includes:

- Number of clients that can be treated
- Experience in specific geographic areas or with specific populations
- Range of services provided
- Rates/cost of treatment.

1.2 Analysis

Contractors receive a quarterly report comparing their performance against a minimum standard as well as statewide averages for the quarterly period, as demonstrated in Exhibit IV-1. Minimum standards are reviewed and revised periodically. Reports are distributed to county alcohol and drug authorities and to individual providers.

Contractors are required to perform above the minimum standard on more than half of the indicators each quarter (e.g., exceed minimum on 5 of 9 indicators for a given provider). Failure to pass minimum standards in one quarter necessitates an action plan; failure to pass minimum standards in three consecutive quarters may result in re-allocation of resources (a reduction in funding).

EXHIBIT IV-1

SAMPLE REPORT FOR A FICTITIOUS OUTPATIENT TREATMENT PROVIDER

Performance Indicator	Minimum Standard (%)	Statewide Average (%)	Provider Average (%)	* Met Std
Employment maintenance	80	93	66.7	
Change in employability	45	65	100.0	*
Employability improvement	15	13	3.0	
Educational advancement	7	33	66.7	*
Participated in self help	20	65	59.5	*
Not arrested during treatment	85	93	97.1	*
Abstinent at termination	37	44	50.0	*
Completed treatment	30	37	42.9	*
Reduced use	40	54	47.6	*

No case-mix adjustments are made to account for potential differences in clients that could explain treatment success. However, the CPMS data system provides information that could be a useful starting point for case-mix, including age, race/ethnicity, gender, previous employment, and nature and extent of substance abuse problems.

Documentation of the performance system does not reference response rates for specific items or entire forms. For example, it is not known whether the discharge form (which drives the performance system) represents all, most, some, or a few clients discharged from treatment in Oregon.

1.3 Impact

Performance data was described as providing a "qualitative" counterpart to the "quantitative" utilization data. State personnel report that no contracts have been terminated solely as a result of this data. They point out that other things often go wrong for contractors that fail to perform relative to standards (for example, utilization may be low). The state may use the performance data to help identify which providers need specific types of technical assistance, and the state may play a role in helping provide that assistance. The state has not undertaken a systematic review of the impact of the performance measurement system on treatment providers.

BEST COPY AVAILABLE

2. METHADONE TREATMENT QUALITY ASSURANCE FEASIBILITY STUDY

Probably the most intensive and advanced effort at evaluating provider performance is the Methadone Treatment Quality Assurance System (MTQAS) Feasibility Study. MTQAS is conducted by the Research Triangle Institute (RTI). After several years of development and planning, MTQAS began a pilot test in seven states in 1996. Since methadone treatment is generally intended to be longer than several months (if not longer than a year), MTQAS only examines the status and performance of current clients.

This is a multi-year research effort sponsored by NIDA to design, test, and operate a performance measurement and reporting system for methadone treatment services (personal communication, William Luckey, Ph.D., 1997). The motivation for MTQAS was to develop an outcome-based accountability system that could supplement or supplant the existing regulatory structure of methadone treatment services (Institute of Medicine, 1995). Current regulation and credentialing of methadone treatment—like other drug treatment, and much of health care—primarily relies on reviews of written policies and procedures, credentialing and training of staff, and selective reviews of case records.

During Phase 1 of the study, RTI developed data collection instruments, assessed their validity and reliability, and pilot tested these instruments with methadone treatment providers. This study is currently in the second phase of activity. During Phase 2, the MTQAS system is being operated in seven states with licensed methadone treatment providers that volunteered to participate. The states that are participating are: Arizona, Colorado, Georgia, Massachusetts, North Carolina, Pennsylvania, and Washington. Approximately 60 percent of the methadone treatment providers within those states have agreed to participate to date, and more are expected to join the effort.

The MTQAS design includes the following sequence of activities:

- Client interviews are conducted at intake. Treatment provider staff conduct face-to-face interviews with clients using a structured interview protocol. All staff are trained by the investigators, through a train-the-trainers process, to administer the survey instrument and collect data in a standardized manner.

- Follow-up interviews plus record abstractions are conducted once every 3 months for clients who have been in treatment for up to a year and once every 6 months for clients enrolled longer than 1 year.

- Records of clients are abstracted again at treatment exit/discharge. Clients are considered discharged if they do not come to the clinical site to receive methadone within a two-week period or if there is an agreement between the client and clinician that methadone treatment will be terminated. Those clients in the former discharge category are scored as being positive for drug use by MTQAS. Those in the latter category are considered as having successfully completed treatment.

Data are submitted to, and processed by, state agencies prior to submission to RTI for analysis. Provider-specific reports are issued to participating providers each quarter for six quarters. These reports provide comprehensive summary information about the status of clients enrolled within each provider, as well as comparisons with other providers within the state and with all providers participating in Phase 2 of the study. Specific measures reported are:

- Urinalysis results for opiates, cocaine, and methadone
- Self-reported use of opiates or cocaine
- Self-reported injection drug use
- Employment status
- Arrests
- Retention
- Self-reported indications that the client has suicidal ideations
- Self-reported health status
- Self-reported health care utilization
- Client satisfaction with treatment.

MTQAS is still very early in development of provider performance measures, and is only now developing case-mix adjusted values for providers. RTI recently issued reports for the second of the six quarters of the pilot test period (RTI, 1996). Reports are sent to the cognizant state agencies as well as the individual providers that participate in MTQAS. For each of the measures indicated above, a value is calculated for each service delivery unit, for all providers in each of the seven states participating in MTQAS, and for all providers across the nation participating in MTQAS. Reports sent to providers include values for their service delivery unit, average values for the state, and the average for the MTQAS providers in the seven states.

A preliminary attempt has been made to adjust provider performance measures for their case mix. As participating providers gain confidence in the data collection requirements, including sampling, data transcription, and reporting, it can be anticipated that the system will be pushed to develop more advanced—and appropriate—measures of provider performance.

V. SUMMARY

The rigorous assessment of provider performance will be increasingly important to the substance abuse treatment field in the future. Provider performance/case-mix analysis will be important and should be useful to:

- Individual providers attempting to assess and improve their own quality of care
- CSAT Knowledge Development and Application initiatives, as well as Cross-site Demonstrations
- Researchers working to learn how to make substance abuse treatment more effective
- Government agencies funding and managing systems of substance abuse treatment
- Managed behavioral health organizations (whether in the private sector, or Medicaid and the public sector)
- Policy makers, Congress, and other government legislative bodies
- Clients and their friends and families.

There is increasing momentum to perform such assessments in the private sector as well as in the public sector. Individual providers (public and private sector) are challenged to develop and implement systems to assess and monitor their own performance in order to maintain quality of care. Provider performance analysis may give providers more systematic data that can be used to assess their own performance, to identify and target potential areas of weakness (and strength), and to improve the quality of care in general and for particular aspects of their service.

Policy makers have put increased importance on evaluating which treatment approach(es) work best for what type of client. Provider performance measurement is a critical element of accountability within the public policy arena. It will be more than an article of good faith to execute such studies and to use the results to manage treatment systems. Pushing such initiatives forward will demonstrate that the principles of efficiency and economy can be rigorously applied to the public sector in its management of important human services. Providing policy makers with the results of performance studies—combined with demonstrated action on the results—will address major concerns that have been expressed for many years about publicly financed substance abuse treatment.

A number of current CSAT Knowledge Development and Application initiatives could apply provider performance methods to good effect. These methods are valuable for evaluation of demonstrations as well as for provider self-monitoring and for management of treatment systems. In any situation where functionally identical data is being acquired from representative samples of clients served by various providers, it is often possible to make provider performance comparisons, which can be used in conjunction with process evaluations to learn more about how well different providers work, and how effective different types of care are.

Several states have operating assessment systems (e.g., Oregon, Maine), and others are developing systems (e.g., Texas, Florida). While it appears that case-mix adjustment is presently not being used in these systems, it should be proposed that this option be explored. The MTQAS for methadone providers represents the most full-blown initiative to perform provider performance analysis (with or without case-mix adjustment) in the substance abuse treatment field to date.

Within the private sector, client outcomes and provider performance are firmly included in the National Committee for Quality Assurance (NCQA, 1996) accreditation standards for managed behavioral health care organizations (MBHOs). Everyone in the field hopes these voluntary standards for MBHOs, developed under the auspices of the industry association by a committee of clinical and managerial experts, can improve the nature and quality (and perhaps more fundamentally, the image) of the industry by documenting procedural protections for and from data about services delivered to persons covered under MBHOs.

Of particular salience for the present discussion is that the NCQA standards require that a MBHO must have a system in place to monitor the performance of providers. Unfortunately, there are no specific requirements for what and how performance should be monitored. However, the NCQA standards represent a movement toward widespread implementation of provider performance assessment. It remains to be seen whether the NCQA will develop or recommend more specific approaches and standards for examining provider performance, or whether the approaches taken by MBHOs will be subjected to public scrutiny and rigorous evaluations.

Within the general health field the hospital mortality rate analyses by the Health Care Financing Administration (HCFA) were for all practical purposes the beginning of the ever more rapidly expanding drive to assess provider performance via client outcomes (Daley et al., 1988; Krakauer et al., 1992 and 1995; Jencks et al., 1988; and Hartz et al., 1988). HCFA initially

developed the estimates as part of intramural research into the feasibility of such measures. These were performed for about 8 years, and discontinued after developing estimates for 1993. Major leaps forward were made in the nature and quality of the methodology over this time period, primarily directed at making adjustments in rankings for particular hospitals based on the nature and severity of their client "case mix."

The "science" of making case-mix adjusted comparisons of provider performance has advanced to the point where there is a fairly clear road map for the substance abuse treatment field. The general approach and methodology is fairly well mapped out. The challenge is to apply this technology to substance abuse treatment. This is not to say that there is little developmental work to do. On the contrary, the existing research literature simply identifies the nature (but not the amount) of work to be done. It remains for CSAT and the substance abuse treatment field to actually undertake the development and testing work that will be necessary in order to realize the promise of this technology for improving the quality and delivery of substance abuse treatment services.

VI. Conclusion

This document presents a discussion of the importance of performance measurement and analysis to the substance abuse treatment field. It also provides tools and processes for measuring treatment outcomes and ensuring the ability to compare outcome measures across programs or against a standard. Data obtained from performance measurement efforts can be used by individual providers to assess and monitor their performance and compare their performance with the performance of other providers. Performance measurement is an integral component of CSAT's approach to evaluation. The goal of any performance measurement strategy should be to generate data that can be used for results-oriented management; provide external accountability; and ultimately improve the quality of services offered. Fundamental to CSAT's approach to evaluation is that evaluation should be an integral part of treatment services management and an ongoing process of assessment and improvement of systems and services. Full implementation and utilization of the Self-Adjusting Treatment Evaluation Model requires incorporating performance measurement efforts with the use of cost analysis, the building of team capabilities, and the assimilation of integrative methodologies.

Performance Measurement for Substance Abuse Treatment has provided:

- A discussion of the increasing importance of provider performance measurement and analysis in the substance abuse field
- A step-by-step explanation of the case-mix adjustment methodology
- A historical perspective on the accomplishments of the substance abuse field in regards to performance measurement.

It is hoped that this information will enable individual service providers and the field to better ensure continuous knowledge development and application.

REFERENCES

Anglin, M. D., Danila, B., Ryan, T., & Mantius, K. (1996). Staying in touch: A fieldwork manual of tracking procedures for locating substance abusers for follow-up studies. Los Angeles, CA: UCLA Drug Abuse Research Center.

Daley, J., Jencks, S. F., Draper, D., Thomas, N., Lenhart, G., & Walker, J. (1998). Predicting hospital-associated mortality for Medicare patients: A method for patients with strokes. Journal of the American Medical Association, 260, 3617-3624.

Devine, P., Christopherson, E., Bishop, S., Lowery, J., Moore, M., et al. (revised 1999). The self-adjusting treatment evaluation model. Fairfax, VA: National Evaluation Data and Technical Assistance Center.

Hartz, A. J., Krakauer, H., Kuhn, E. M., Young, M., Jacobsen, S. J., Gay, G., Muenz, L., Katzoff, M., Bailey, R. C., & Rimm, A. A. (1989). Hospital characteristics and mortality rates. New England Journal of Medicine, 321, 1720-1724.

Harwood, H. (revised 1999). Adding value to CSAT demonstrations: The what, how, and why of cost analysis. Fairfax, VA: National Evaluation Data and Technical Assistance Center.

Institute of Medicine (1995). Regulating methadone treatment. Washington, DC: National Academy Press.

Jencks, S. F., Daley, J., Draper, D., Thomas, N., Lenhart, G., & Walker, J. (1989). Interpreting hospital mortality data: The role of clinical risk adjustment. Journal of the American Medical Association, 260, 3611-3616.

Kahn, K. L., Brook, R. H., Draper, D., Keeler, E. B., Rubenstein, L.V., Rogers, W. H., & Kosecoff, J. (1988). Interpreting hospital mortality data: How can we proceed? Journal of the American Medical Association, 260, 3625-3628.

Krakauer, H., Bailey, R. C., Cooper, H., Yu, W. K., Skellan, K. J., & Kattakkuzhy, G. (1995, January-February). The systematic assessment of variations in medical practices and their outcomes. Public Health Reports, 110, 2-12.

Krakauer, H., Bailey, R. C., Skellan, K. J., Stewart, J. D., Hartz, A. J., Kuhn, E. M., & Rimm, A. A. (1992). Evaluation of the HCFA model for the analysis of mortality following hospitalization. Health Services Research, 27, 317-335.

McLellan, A., Grissom, G., Brill, P., Durell, J., Metzger, D., & O'Brien, C. (1993). Private substance abuse treatments: Are some programs more effective than others? Journal of Substance Abuse Treatment, 3, 243-245.

Moore, M. (revised 1999). Building team capability to fully implement and utilize the self-adjusting treatment evaluation model. Fairfax, VA: National Evaluation Data and Technical Assistance Center.

National Committee For Quality Assurance (1996). Standards for accreditation of managed behavioral health care organizations, 1997. Annapolis Junction, MD: Author.

National Evaluation Data and Technical Assistance Center. (1999). Training for CSAT staff and grantees on conducting cross-site evaluation. Final Edition: Volume I. Fairfax, VA.

Office of Applied Studies (1995). Drug and alcohol services information system state manual. Rockville, MD: Substance Abuse and Mental Health Services Administration.

Oregon Office of Alcohol and Drug Abuse Programs (1996). Performance indicators. Salem, Oregon: Author.

Oregon Office of Medical Assistance Programs (1996). Chemical dependency contract standards in the Oregon health plan: Fully capitated health plan agreement. Salem, Oregon: Author.

Phibbs, C., Swindle, R., & Recine, B. (1997). Does case mix matter for substance abuse treatment? A comparison of observed and case mix-adjusted readmission rates for inpatient substance abuse treatment in the Department of Veteran Affairs. Health Services Research, 6, 755-771.

Research Triangle Institute (1996). Quarterly feedback report: Methadone quality assurance system (Draft). Research Triangle Park, NC: Author.

APPENDIX A:
INTEGRATED EVALUATION METHODS PACKAGE:
A GUIDE FOR SUBSTANCE ABUSE TREATMENT
KNOWLEDGE-GENERATING ACTIVITIES—EXECUTIVE SUMMARY

APPENDIX A:
INTEGRATED EVALUATION METHODS PACKAGE:
A GUIDE FOR SUBSTANCE ABUSE TREATMENT
KNOWLEDGE-GENERATING ACTIVITIES—EXECUTIVE SUMMARY

Since its inception, the Center for Substance Abuse Treatment (CSAT) has provided Federal leadership to improve substance abuse treatment accessibility, effectiveness, and efficiency. CSAT's mission and activities have evolved from directly supporting treatment services to supporting knowledge-generating activities. This evolution is evident in the current Substance Abuse and Mental Health Services Administration policy on evaluation as described in *Evaluation Policy*, SAMHSA, 1995.

The need for an integrated model of evaluation and planning at SAMHSA is presented in "Evaluation in the Substance Abuse and Mental Health Services Administration," *Evaluation and the Health Professions*, by Marsh, Jansen, Lewis, & Straw, 1996. CSAT also supports site-specific, cross-site, and national evaluations that have provided experience with a wide array of evaluation design and implementation methods. These experiences further supported the need for an integrated evaluation strategy and led to the development of a comprehensive set of evaluation products, including concept papers, technical assistance (TA) materials, and analytic tools. Collectively, these products are referred to as the Integrated Evaluation Methods (IEM) Package. The IEM Package organizes these products within an evaluation framework that is designed to support CSAT knowledge development and application goals. The evaluation framework itself was constructed on the basis of accumulated experiences among internationally known treatment service evaluation professionals. The IEM Package reflects and incorporates evaluation experiences gained over the past decade.

Evaluation Framework and the Integrated Evaluation Methods Package

National evaluation experiences have reinforced the fact that substance abuse treatment evaluation involves a standard set of tasks that generally occur in the following order:

- **Planning the evaluation/knowledge-generating activities**, which includes selecting the substance abuse treatment issue, identifying the theoretical foundation for the intervention, determining knowledge development program goals and implementation approach, and setting the evaluation goals and objectives that determine the overall parameters of the evaluation

- **Selecting the evaluation design**, which sets forth the overall strategy for establishing the process and outcome evaluation questions, measurement approach, and generalizability of findings

- **Developing the data requirements**, which flow from the evaluation questions and measures and include: SDU, clinician, cost, and client data

- **Developing data collection instruments**, which are based on the data requirements and are developed or selected from an integrated inventory of instrumentation

- **Collecting the data**, which includes developing data management processes and tools (including quality control procedures) and conducting the data collection activities

- **Analyzing the data**, which involves multiple levels of comparison and is governed by an analysis plan

- **Reporting the evaluation findings**, which includes evaluation knowledge dissemination and application within the field.

The evaluation process outlined above provided a framework for the development of products related to these evaluation concepts and methods. Taken together, those products comprise the IEM Package.

Integrated Evaluation Methods Products

CSAT requested the development of a series of evaluation concept papers, TA materials, and tools to support and operationalize each phase in the evaluation of substance abuse treatment knowledge-generating activities. These items are included in the IEM Package. The concept papers are based on theoretical evaluation research constructs that have been adapted to substance abuse treatment services evaluation and knowledge-generating activities. The concept papers primarily support the evaluation planning phase and address such topics as the self-adjusting treatment evaluation model, cost analyses, and performance measurement. The TA materials and tools include specific evaluation methods that have direct applicability to substance abuse treatment knowledge-generating activities. The concept papers and TA materials that constitute the IEM Package are listed and briefly described in Exhibit I.

EXHIBIT I
EVALUATION FRAMEWORK AND INTEGRATED EVALUATION METHODS PACKAGE

EVALUATION FRAMEWORK	INTEGRATED EVALUATION METHODS PRODUCTS
1. Planning the evaluation/ knowledge-generating activities	■ **Integrated Evaluation Methods: A Guide for Substance Abuse Treatment Knowledge Generating Activities:** Concept paper that describes the development of an evaluation framework, evaluation concepts, and TA materials to support the framework. ■ **Self-Adjusting Treatment Evaluation Model:** Concept paper that describes an approach for integrating evaluation findings within treatment operations so as to adjust and improve service delivery. ■ **Building Team Capability to Fully Implement and Utilize the Self-Adjusting Treatment Evaluation Model:** Concept paper to assist treatment providers in building capabilities to integrate the self-adjusting treatment model within day-to-day operations and service delivery. ■ **Adding "Value" to CSAT Demonstrations:** The What, How and Why of Cost Analysis: Concept paper on the need for and types of cost analyses for CSAT demonstrations and knowledge-generating activities. (The Lewin Group) ■ **Performance Measurement for Substance Abuse Treatment Services:** Concept paper about the increasing importance of provider performance measurement and analyses and an explanation of the case-mix adjustment methodology. ■ **Client Levels of Functioning as a Component of Substance Abuse Treatment Services Evaluation**: Description of the rationale and methods for assessing client level of functioning and recommended core LOF data elements that could help to measure the effectiveness of treatment services received. ■ **Substance Abuse Treatment Evaluation Policy Notebook:** These materials are aimed at facilitating understanding of the SAMHSA policy for evaluation and federal regulations on client confidentiality and assisting evaluators to meet CSAT evaluation requirements. ■ **Substance Abuse Treatment Evaluation Resource Notebook:** The notebook contains evaluation bibliographies and listings of organizations, hot lines, on-line data bases, and contact information for obtaining assistance in evaluating treatment services.
2. Selecting the evaluation design	■ **A Guide to Process Evaluation for Substance Abuse Treatment Services:** TA tool presenting purposes of process evaluation and the application of process evaluation methods to single site and multi-site treatment services. ■ **Using Logic Models in Substance Abuse Treatment Evaluations:** TA tool describing logic model purposes and techniques for designing and planning the evaluation of treatment services. ■ **A Guide to Selecting an Outcome Evaluation Design for Substance Abuse Treatment Evaluations:** TA tool describing overall strategies for developing evaluation questions, measurement, controls, validity/reliability, sampling, design effects, and generalizability of findings. (Battelle)

BEST COPY AVAILABLE

EXHIBIT I (CONTINUED)
EVALUATION FRAMEWORK AND INTEGRATED EVALUATION METHODS PACKAGE

EVALUATION FRAMEWORK	INTEGRATED EVALUATION METHODS PACKAGE
3. Developing data requirements	■ **Minimum Evaluation Data Set (MEDS): Core Data Lists:** TA tool for developing a uniform set of variables and response categories for the service delivery unit (SDU), clinician, cost, and client evaluation measures. ■ **Substance Abuse Treatment Cost Allocation and Analysis Template (SATCAAT):** User manual to analyze treatment costs by unit of service for an SDU. (Capital Consulting Corporation)
4. Developing data collection instruments	■ **Substance Abuse Treatment Services Evaluation Data Collection Instruments:** Data collection instruments that fully incorporate the MEDS and that have been field tested for validity and reliability, as follows: Service Delivery Unit (SDU) Description; Clinician Background and Practice Survey; protocols to collect Adult, Adolescent and Child (in treatment with parent) Client Data at Intake, During Treatment, at Treatment Discharge and Post Treatment; Adult and Adolescent Record Extraction forms; and a section on protection of human subjects and informed consent.
5. Collecting the data	■ **Staying In Touch: A Fieldwork Manual of Tracking Procedures for Locating Substance Abusers for Follow-up Studies (UCLA):** User manual to establish and implement client follow-up data collection systems and procedures. ■ **Strategies for Follow-up Tracking of Juvenile, Homeless, and Criminal Justice System-Involved Substance Abusers: Overview and Bibliographies, 1990-1998:** Description of tracking techniques used to increase response rates for follow-up interviews with homeless and juvenile/criminal justice involved substance abusers.
6. Analyzing the data	■ **A Guide to Substance Abuse Treatment Evaluation Data Analysis:** Recommended methods and procedures for analyzing process, SDU, clinician, cost, and client evaluation data.
7. Reporting the evaluation findings	■ **Substance Abuse Treatment Evaluation Product Outlines Notebook:** Compendium of outlines for evaluation products including evaluation plans, interim evaluation reports, final evaluation reports, replication studies, case studies, and ethnographies.

BEST COPY AVAILABLE

CSAT Evaluation "Stakeholders"

Evaluation "stakeholders" are individuals, groups, or organizations that have a significant interest in how well a program or activity functions. (See P.H. Rossi, H.E. Freeman, & M.W. Lipsey, *Evaluation: A Systematic Approach, 6th Edition*, 1999.) Within the context of the IEM Package, CSAT evaluation stakeholders include CSAT senior managers, CSAT project officers, and CSAT grantees and contractors including treatment service providers, coordinating centers, study sites, site-specific evaluators, and national evaluators.

Utility of the IEM Package for CSAT Evaluation Stakeholders

While the conceptual and TA materials were developed from the perspective of the site-specific and multi-site evaluator, the concepts and TA tools have important utility for CSAT managers, project officers, and treatment service providers. The stakeholder's position determines the perspective and utility of the IEM Package concepts and tools. For example, a CSAT senior manager can use the IEM Package to acquire a comprehensive evaluation context for planning and funding the knowledge-generating activities, the project officer can use the IEM Package to ensure that GFA/RFP applications are complete and include a full complement of design, execution, and product components, and the site-specific and multi-site evaluators can use the IEM Package to ensure that evaluation designs, data collection plans, data analyses, and product development have a consistent evaluation framework and compatible data across program areas. The suggested utility of the IEM Package for CSAT evaluation stakeholders is summarized in Exhibit II.

EXHIBIT II
UTILITY OF IEM PACKAGE FOR CSAT EVALUATION STAKEHOLDERS

STAKEHOLDERS	ROLES AND RESPONSIBILITIES	IEM PACKAGE UTILITY
SENIOR MANAGERS	■ Policy development ■ Issue identification for KD&As ■ Grant/contract funding decisions ■ Overall program management ■ Sustainability ■ Dissemination ■ Long-term strategic planning ■ Program designs ■ KA activities	■ Comprehensive evaluation framework ■ Comprehensive evaluation components ■ Roles and responsibilities for local/national evaluators as well as CSAT/grantee staffs ■ Guidance for evaluation designs and products ■ Standardized evaluation measures ■ Logic models for program and evaluation design
PROJECT OFFICERS	■ GFA/SOW development ■ Grant/contract application review ■ Grant/contract monitoring ■ Knowledge-generating products ■ Identification and replication of promising practices ■ Technical assistance assessment	■ Guidelines for high-quality evaluation designs (process and outcome) ■ Logic models for program and evaluation designs ■ List of evaluation measures with instrumentation ■ Guidelines for evaluation products
GRANTEES: STUDY SITES	■ Grant applications ■ Project development, implementation ■ Local evaluation management ■ Local evaluation coordination ■ Knowledge-generating product development	■ Evaluation plan outline ■ Process and outcomes evaluation designs ■ SDU, clinician, cost, and client measures ■ Roles and responsibilities for grantee provider/evaluator staff ■ Guidelines for evaluation products
GRANTEES: MULTI-SITE EVALUATORS	■ Grant applications ■ Comprehensive evaluation designs ■ Evaluation implementation: – Data collection – Data analysis – Reporting evaluation findings ■ Evaluation product development	■ Evaluation concepts ■ Logic models ■ Evaluation designs ■ Evaluation data requirements ■ Data collection instrumentation ■ Data collection process and procedures ■ Data analysis ■ Product development
NATIONAL EVALUATORS/ SERVICES RESEARCHERS	■ Contract applications ■ Comprehensive evaluation designs ■ Evaluation implementation: – Data collection – Data analysis – Reporting evaluation findings ■ Evaluation product development	■ Evaluation concepts ■ Logic models ■ Evaluation designs ■ Evaluation data requirements ■ Data collection instrumentation ■ Data collection process and procedures ■ Data analysis ■ Product development

IEM products and other evaluation materials may be obtained from:
http://neds.calib.com

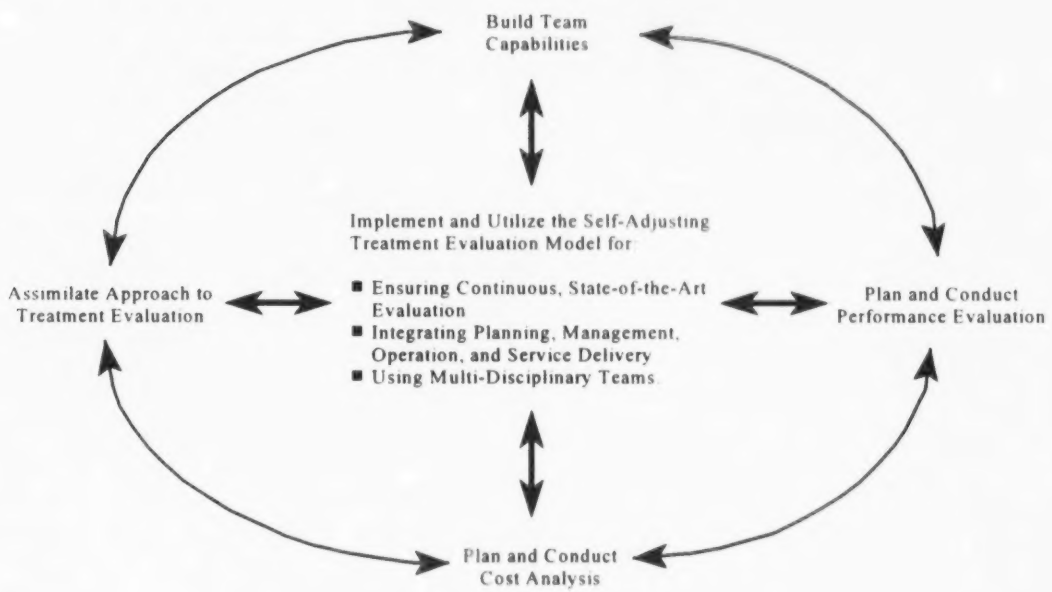
BEST COPY AVAILABLE

APPENDIX B:
EDITOR'S NOTE

EDITOR'S NOTE

This document is one of a series of papers that describe CSAT's approach to substance abuse treatment evaluation. The graphic below illustrates the continuous evaluation knowledge development and application process which characterizes CSAT's approach. At the core is the self-adjusting treatment evaluation model which is the foundation. The model integrates continuous, state-of-the-art evaluation with planning, management, operation, and service delivery within a multi-disciplinary learning community. Implementation of this model requires building of team capabilities, appropriate, state-of-the-art performance evaluation and cost analysis, and assimilation of CSAT's integrative approach to treatment evaluation and integrative methodologies. Each of these processes work together to ensure continuous improvement.

ENSURING CONTINUOUS EVALUATION KNOWLEDGE DEVELOPMENT AND APPLICATION


Build Team Capabilities
Implement and Utilize the Self-Adjusting Treatment Evaluation Model for:
Ensuring Continuous, State-of-the-Art Evaluation
Integrating Planning, Management, Operation, and Service Delivery
Using Multi-Disciplinary Teams
Assimilate Approach to Treatment Evaluation
Plan and Conduct Performance Evaluation
Plan and Conduct Cost Analysis

Substance abuse treatment providers are increasingly called upon to demonstrate that they are delivering appropriate services, that those services have the desired impact, and that the services justify the costs. An ongoing process of evaluation and systems/services improvement integrated into the day-to-day operation of treatment providers is needed to do so. In addition, the evaluation and improvement process requires a multi-disciplinary team that includes treatment personnel, evaluators, Federal and state agencies, advocacy groups, funding agencies, and the community. Building team capability is integral to this approach. Treatment staff must be involved in knowledge development and application (i.e., planning and implementing evaluation efforts, incorporating changes in response to new knowledge, and sharing of findings).